[Title of Document]	Application for Patent
[Reference Number]	9800629303
[Date of Filing]	August 24, 1998
[Addressee]	Director-General, Patent Office Takeshi ISAYAMA
[International Patent Classification]	G03B 17/18
[Title of the Invention]	Encoding Apparatus and Digital Camera Apparatus
[Number of Claim(s)]	12
[Inventor]	
[Domicile or Residence]	c/o SONY CORPORATION 7-35, Kitashinagawa 6-chome, Shinagawa-ku, Tokyo, Japan
[Name]	Shigeki ISHIZUKA
[Inventor]	
[Domicile or Residence]	c/o SONY CORPORATION 7-35, Kitashinagawa 6-chome, Shinagawa-ku, Tokyo, Japan
[Name]	Katsuhiko UENO
[Inventor]	
[Domicile or Residence]	c/o SONY CORPORATION 7-35, Kitashinagawa 6-chome, Shinagawa-ku, Tokyo, Japan
[Name]	Yoshihiro ONO
[Applicant]	

【ID Number】	000002185
【Name】	SONY CORPORATION
【Representative】	Nobuyuki IDEI
【Agent】	
【ID number】	100082762
【Patent Attorney】	
【Name】	Masatomo SUGIURA
【Telephone number】	03-3980-0339
【Indication of Fee】	
【Prepayment Register Number】	043812
【Amount of Payment】	¥21,000
【List of Items Filed】	
【Title of Article】	Specification 1
【Title of Article】	Drawings 1
【Title of Article】	Abstract 1
【Necessity of Proof】	Yes

[Title of Document] Specification

[Title of the Invention] ENCODING APPARATUS AND DIGITAL CAMERA APPARATUS

[Scope of Claims for a Patent]

[Claim 1]

An encoding apparatus for encoding a video signal in MPEG video format, encoding an audio signal in MPEG audio format, multiplexing the encoded video signal as MPEG video data and the encoded audio signal as MPEG audio data, and generating the multiplexed data, comprising:

vide encoding means for encoding still picture data corresponding to intra-frame encoding process so as to generate an I picture, generating a P picture or a B picture in such a manner that moving vectors of all macro blocks thereof are zero and the chronologically preceding picture is copied as an encoded picture, and generating an MPEG video data in a frame structure of which the I picture is followed by a plurality of P pictures or B pictures,

wherein the time period of the P pictures or the B pictures is almost the same as the time period of the audio signal encoded in the MPEG audio format.

[Claim 2]

The encoding apparatus as set forth in claim 1,

wherein the multiplexed data is recorded to a

storing medium.

[Claim 3]

The encoding apparatus as set forth in claim 1,

wherein the multiplexed data is transmitted to a communication path.

[Claim 4]

A digital camera apparatus for recording a photographed picture as a digital signal to a record medium, comprising:

photographing means for outputting a photographed still picture;

signal processing means for processing a signal received from said photographing means;

video encoding means for encoding a digital picture signal received from said signal processing means in MPEG format and generating MPEG video data;

audio inputting means;

audio encoding means for converting an input audio signal into a digital audio signal, encoding the digital audio signal in MPEG audio format, and generating MPEG audio data;

memory means for storing multiplexed data of the MPEG video data and the MPEG audio data;

controlling means for controlling a storing operation of the multiplexed data to said memory means;

displaying means for displaying the digital

picture signal;

a storing medium and storing medium driving means for storing the multiplexed data stored in said memory means; and

operating means including a shutter button,

wherein said video encoding means encodes the photographed still picture data corresponding to intra-frame encoding method so as to generate an I picture, generates a P picture or a B picture in such a manner that moving vectors of all macro blocks thereof are zero and the chronologically preceding picture is copied as an encoded picture, and outputs a video encoded signal in a frame structure of which the I picture is followed by a plurality of P pictures or B pictures.

[Claim 5]

The digital camera apparatus as set forth in claim 4,

wherein said audio encoding means encodes an audio signal after a still picture is photographed until a predetermined time period elapses and generates the resultant signal as MPEG audio data.

[Claim 6]

The digital camera apparatus as set forth in claim 4,

wherein when the multiplexed data is stored to said memory means, said controlling means controls

said memory means and said storing medium driving means so as to read the multiplexed data from said memory means and record the multiplexed data to said storing medium.

[Claim 7]

The digital camera apparatus as set forth in claim 4, further comprising:

video decoding means for decoding the MPEG video data;

audio decoding means for decoding the MPEG audio data; and

audio reproducing means,

wherein said controlling means controls said memory means and said storing medium driving means so as to reproduce the multiplexed data from said storing medium and store the reproduced multiplexed data to said memory means, and

wherein the MPEG video data received from said memory means is decoded by said video decoding means, the decoded picture data is displayed on said displaying means, the MPEG audio data received from said memory means is decoded by said audio decoding means, and the decoded audio data is reproduced by said audio reproducing means.

[Claim 8]

The digital camera apparatus as set forth in claim 4,

wherein the multiplexed data is a stream composed of a plurality of packs, the MPEG audio data and the I picture of the MPEG video data being placed at the top pack.

[Claim 9]

A digital camera apparatus for recording a photographed picture as a digital signal to a record medium, comprising:

photographing means for outputting a photographed still picture;

signal processing means for processing a signal received from said photographing means;

first video encoding means for encoding a digital picture signal received from said signal processing means and generating first encoded video data;

second video encoding means for encoding a digital picture signal received from said signal processing means and generating second encoded video data;

audio inputting means;

audio encoding means for converting an input audio signal into a digital audio signal, encoding the digital audio signal, and generating encoded audio data;

controlling means for controlling a storing operation of data to memory means;

displaying means for displaying the digital picture signal;

a storing medium and storing medium driving means for storing data stored in said memory means; and

operating means including a shutter button,

wherein an output signal of the first encoded video data and an output signal of which the second encoded video data and the encoded audio data are multiplexed.

[Claim 10]

The digital camera apparatus as set forth in claim 9,

wherein said first video encoding means generates the first encoded video data in JPEG format,

wherein said second video encoding means generates the second encoded video data in MPEG format, and

wherein said audio encoding means generates the encoded audio data in MPEG audio format.

[Claim 11]

The digital camera apparatus as set forth in claim 9,

wherein said controlling means controls a first process for writing the digital picture signal to the memory means, a second process for writing multiplexed data of the first encoded video data and the encoded audio data to the memory means, and a third

process for reading the digital picture signal from the memory means, supplying the digital picture signal to said second video encoding means, and writing the second encoded video data to the memory means.

[Claim 12]

The digital camera apparatus as set forth in claim 10,

wherein after the multiplexed data of the first encoded video data and the encoded audio data is written to the memory means, the multiplexed data is read from the memory means and then stored to said storing medium, after the multiplexed data is stored, said second video encoding means encodes the digital picture signal and generates the second encoded video data, after the second encoded video data is written to the memory means, the second encoded video data is read from the memory means and then stored to said storing medium.

[Detailed Description of the Invention]

[0001]

[Technical Field to which the Invention belongs]

The present invention relates to an encoding apparatus applicable for a digital still camera that records a photographed still picture to a record medium. The present invention also relates to such a digital camera apparatus.

[0002]

[Prior Art]

Digital cameras that record digital picture information to record mediums such as a floppy disk and a semiconductor memory are becoming common. A digital camera converts a photographed picture into a digital picture signal, compresses the digital picture signal, and records the compressed picture information to a record medium. A digital camera can also record a moving picture as well as a still picture.

[0003]

JPEG (Joint Photographic Experts Group) format that is a format for compressing a still picture and MPEG (Moving Picture Experts Group) format are general-purpose encoding formats adopted by ISO. These formats are suitable for picture data photographed by a digital camera and used in a personal computer. In the JPEG format, a color still picture is compression-encoded by DCT (Discrete Cosine Transform) method. Coefficient data is quantized. The quantized output is encoded with variable length code. In contrast, in the MPEG format, a color moving picture is compression-encoded and a frame difference between an input picture and a motion-compensated predictive picture is encoded by DCT method.

[0004]

[Subject that the Invention is to solve]

When a digital camera can record a still

picture and an audio signal corresponding thereto, a memo of a still picture can be recorded as an audio signal. However, since the JPEG format is designed to record and transmit information of still pictures, audio information corresponding to still pictures , cannot be simultaneously recorded and transmitted. Likewise, in other still picture formats (GIF, TIFF, BMP, and so forth), a still picture and audio information corresponding thereto cannot be simultaneously recorded and transmitted. Although software that allows a still picture and audio information corresponding thereto to be simultaneously recorded and transmitted is known (for example, Exif), it is not common. Even if audio attached still picture data is created using such software, software for a player that reproduces the audio attached still picture data is not easily available.

[0005]

Therefore, an object of the present invention is to provide an encoding apparatus that encodes a still picture and audio information corresponding thereto in MPEG format that is a general-purpose format.

[0006]

Another object of the present invention is to provide an encoding apparatus and a digital camera apparatus that simultaneously record a photographed

still picture and audio information corresponding thereto.

[0007]

To solve such a problem, the invention of claim 1 is an encoding apparatus for encoding a video signal in MPEG video format, encoding an audio signal in MPEG audio format, multiplexing the encoded video signal as MPEG video data and the encoded audio signal as MPEG audio data, and generating the multiplexed data, comprising a vide encoding means for encoding still picture data corresponding to intra-frame encoding process so as to generate an I picture, generating a P picture or a B picture in such a manner that moving vectors of all macro blocks thereof are zero and the chronologically preceding picture is copied as an encoded picture, and generating an MPEG video data in a frame structure of which the I picture is followed by a plurality of P pictures or B pictures, wherein the time period of the P pictures or the B pictures is almost the same as the time period of the audio signal encoded in the MPEG audio format.

[0008]

The invention of claim 4 is a digital camera apparatus for recording a photographed picture as a digital signal to a record medium, comprising a photographing means for outputting a photographed still picture, a signal processing means for processing a

signal received from the photographing means, a video encoding means for encoding a digital picture signal received from the signal processing means in MPEG format and generating MPEG video data, an audio inputting means, an audio encoding means for converting an input audio signal into a digital audio signal, encoding the digital audio signal in MPEG audio format, and generating MPEG audio data, a memory means for storing multiplexed data of the MPEG video data and the MPEG audio data, a controlling means for controlling a storing operation of the multiplexed data to the memory means, a displaying means for displaying the digital picture signal, a storing medium and a storing medium driving means for storing the multiplexed data stored in the memory means, and a operating means including a shutter button, wherein the video encoding means encodes the photographed still picture data corresponding to intra-frame encoding method so as to generate an I picture, generates a P picture or a B picture in such a manner that moving vectors of all macro blocks thereof are zero and the chronologically preceding picture is copied as an encoded picture, and outputs a video encoded signal in a frame structure of which the I picture is followed by a plurality of P pictures or B pictures.

[0009]

The invention of claim 9 is a digital camera

apparatus for recording a photographed picture as a digital signal to a record medium, comprising a photographing means for outputting a photographed still picture, a signal processing means for processing a signal received from the photographing means, a first video encoding means for encoding a digital picture signal received from the signal processing means and generating first encoded video data, a second video encoding means for encoding a digital picture signal received from the signal processing means and generating second encoded video data, an audio inputting means, an audio encoding means for converting an input audio signal into a digital audio signal, encoding the digital audio signal, and generating encoded audio data, a controlling means for controlling a storing operation of data to memory means, a displaying means for displaying the digital picture signal, a storing medium and a storing medium driving means for storing data stored in the memory means, and an operating means including a shutter button, wherein an output signal of the first encoded video data and an output signal of which the second encoded video data and the encoded audio data are multiplexed.

[0010]

According to the invention of claim 1, when still picture data is recorded or transmitted, audio information corresponding thereto can be multiplexed

with the still picture data. Thus, a still picture and audio information that has been recorded can be reproduced by a personal computer using general purpose software.

[0011]

According to the invention of claim 4, when a still picture is photographed, audio information corresponding thereto can be recorded. The still picture and the audio information can be multiplexed corresponding to the MPEG format. Thus, a still picture and audio information that has been recorded can be reproduced by a personal computer using software that is commercially available.

[0012]

According to the invention of claim 9, a function for simultaneously recording a still picture and an audio signal can be accomplished for a digital camera. In addition, when an audio attached still picture is recorded, only a still picture can be recorded. Thus, corresponding to a desired application, recorded data can be used.

[0013]

[Embodiment of the Invention]

Next, a digital camera according to an embodiment of the present invention will be described. The digital camera according to an embodiment of the present invention can photograph and record a still

picture, an audio attached still picture, and an audio attached moving picture. Fig. 1 shows the overall structure of the digital camera according to the embodiment of the present invention. Referring to Fig. 1, a photographing portion is composed of a lens portion 1 and a CCD (Charge Coupled Device) 2. A control signal is supplied from a CPU 12 to the lens portion 1. In the lens portion 1, an automatic diaphragm control operation and an automatic focus control operation are performed corresponding to the control signal received from the CPU 12. The CCD 2 has a photographing mode and a line thin-out mode (referred to as E-to-E mode). In the photographing mode, all pixels are read. In the line thin-out mode, the number of lines are thinned out by 3. The CC2 selects one of the photographing mode and the line thin-out mode corresponding to a control signal received from the CPU 12. The number of pixels of the CCD 2 is 1024 x 768 corresponding to XGA (extended Graphics Array).

[0014]

Next, the real operation of the CCD 2 will be described. In the still picture photographing mode, signal electric charges are read from photo sensors to a vertical CCD. The signal electric charges of all the pixels are successively transferred to a horizontal CCD. In the E-to-E mode or a moving picture photographing mode (that will be described later),

since the number of lines through which signal electric charges of photo sensors are supplied to transfer gates is divided, the number of lines is thinned out by for example 3.

[0015]

According to the present invention, a solid state image pickup device (not limited to a CCD) that thins out the number of lines in other than the above-described structure, a solid state image pickup device that thins out the number of pixels in horizontal direction, or a solid state image pickup device that thins out the number of lines in vertical direction and the number of pixels in horizontal direction can be used.

[0016]

In the E-to-E mode, a photographed picture is displayed on a displaying portion (LCD 8), not stored in a memory (DRAM 9). In the E-to-E mode, when a picture is photographed, the angle of view, focus point, exposure, and white balance are adjusted. In other words, the state of which the user checks an object before pressing the shutter button is the E-to-E mode. In the E-to-E mode, a photographed signal of 1024 x 256 pixels is obtained from the CCD 2. For example, in the photographing mode, a photographed signal of 10 frames per second is output. In contrast, in the E-to-E mode, a photographed signal of 30 frames

per second is output.

[0017]

An output signal of the CCD 2 is supplied to a sample hold and A/D converting portion 3. The sample hold and A/D converting portion 3 generates a digital photographed signal of 10 bits per sample. The sample and A/D converting portion 3 is composed of a correlative dual sampling circuit so as to remove noise, trim waveform, and compensate defective pixels.

[0018]

The digital photographed signal is supplied to a camera signal processing portion 4. The camera signal processing portion 4 includes a digital clamping circuit, a luminance signal processing circuit, a color signal processing circuit, a contour compensating circuit, a defect compensating circuit, an automatic diaphragm controlling circuit, an automatic focus controlling circuit, an automatic white balance compensating circuit, and so forth. The camera signal processing portion 4 generates a digital component signal (composed of a brightness signal and color difference signals) into which an RGB signal is converted.

[0019]

Components of the digital photographed signal are supplied from the camera signal processing portion 4 to a memory controller 5. The memory controller 5 is

connected to a display buffer memory 6 and a bus 14 of the CPU 12. The buffer memory 6 processes a component signal, generates an RGB signal, and outputs the RGB signal to a D/A converter 7. The D/A converter 7 supplies an analog signal to the LCD 8. The buffer memory 6 outputs the RGB signal at a timing corresponding to a display timing of the LCD 8.

[0020]

The bus 14 is connected to the DRAM (Dynamic Random Access Memory) 9, the CPU 12, an encoder/decoder 15, and an interface 10. The DRAM 9 is controlled corresponding to an address signal and a control signal received from the memory controller 5 and the CPU 12, respectively. The memory controller 5 has a pixel number converting function for converting the number of pixels corresponding to a picture size or an operation mode that are set by the user.

[0021]

For example, as shown in Fig. 2, a picture can be recorded in one of picture formats XGA, VGA (Video Graphics Array: 640 x 480 pixels), CIF (Common Intermediate Format: 320 x 240 pixels), and QCIF (Quater CIF: 160 x 120 pixels). However, since the size of each macro block in the MPEG format is 16 x 16 pixels, a picture in the picture format QCIF is composed of 160 x 112 pixels. In other words, the upper portion and the lower portion of a picture in the

picture format QCIF are removed. In the picture format XGA, a photographed signal of the CCD 2 is directly recorded. The picture formats XGA and VGA are used for still pictures. The picture format CIF is used for an audio attached still picture. The picture formats CIF and QCIF are used for audio attached moving pictures.

[0022]

The encoder/decoder 15 compress (encodes) or decompresses (decodes) picture data. For example, when a still picture is processed, the JPEG (Joint Photographic Experts Group) format is used. For example, when a moving picture is processed, the MPEG (Moving Picture Experts Group) format is used. The encoder/decoder 15 has functions corresponding to both the encoding formats. In reality, as a format for compressing a moving picture, MPEG1 format is used.

[0023]

In the MPEG1 format, there are three picture types that are an I picture, a P picture, and a B picture. When an I picture is encoded, only the information thereof is used. Thus, an I picture can be decoded with only information thereof. When a P picture is encoded, as a predictive picture (that is a reference picture for obtaining a difference), an I picture or P picture that has been decoded chronologically before the current P picture is used. The difference between the current P picture and a

predictive picture that has been motion-compensated is encoded or the current P picture is encoded. One of the encoding processes is selected block by block whichever effective. When a B picture is encoded, an I picture or a P picture that has been decoded as a predictive picture chronologically before the current B picture, an I picture or a P picture that has been decoded as a predictive picture chronologically before the current B picture, and an interpolated picture of these predictive pictures are used. The difference between the current picture and each of the predictive pictures that have been motion-compensated is encoded or the current B picture is encoded. One of the encoding processes is selected block by block whichever the most effective.

[0024]

Thus, there are four types of macro blocks that are an intra-frame encoded macro block, a forward inter-frame predictive macro block of which a future macro block is predicted with a past macro block, a backward inter-frame predictive macro block of which a past macro block is predicted with a future macro block, and a bidirectional inter-frame predictive macro block of which the current macro block is predicted with a future macro block and a past macro block. All macro blocks of an I picture are intra-frame encoded macro blocks. A P picture contains intra-frame encoded

macro blocks and forward inter-frame predictive macro blocks. A B picture contains all the four types of macro blocks.

[0025]

In the MPEG1 format, a DCT process is performed for each block composed of (8x8 pixels). A macro block is composed of four luminance (Y) block and two color difference (Cb and Cr) blocks. A slice layer is composed of a predetermined number of macro blocks. A picture layer is composed of a plurality of slice layers. A macro block layer contains a code that represents a macro block type, a code equivalent to a skip of 33 macro blocks, a code that represents (the number of macro blocks to be skipped plus 1), a horizontal component and a vertical component of a moving vector, a code that represents whether or not the six blocks of the current macro block have coefficients, and so forth. The MPEG1 format defines that the first macro block and the last macro block of a slice cannot be skipped. The slide layer contains a code that represents the beginning of the current slice layer.

[0026]

According to the embodiment of the present invention, when an audio attached still picture or an audio attached moving picture is recorded, video data is encoded in the MPEG format. As will be described

later, the encoder/decoder 15 performs an MPEG encoding process omitting the motion compensation inter-frame predictive process so as to reduce the amount of generated code.

[0027]

The interface 10 is an interface between an external storing medium 11 and the CPU 12. Examples of the external storing medium are a disk type recording medium (such as a floppy disk) and a memory card. An operation signal is supplied from an operation and inputting portion 13 to the CPU 12. The operation and inputting portion 13 includes a shutter button and various switches that the user operates. In addition, the operation and inputting portion 13 includes a photographing (recording) mode switch of the digital camera and a picture size switch for designating the size of a picture stored to the external storing medium. The operation and inputting portion 13 detects an operation of each button and each switch and supplies the detected signal as an operation signal to the CPU 12. The shutter speed and the diaphragm are automatically set corresponding to an object and a photographing condition. The digital camera may have a plurality of photographing modes as well as the automatic mode.

[0028]

When a picture is photographed by the digital

camera, the CCD 2 is set to the E-to-E mode. The angle of view, focus, and exposure are properly set. In the E-to-E mode, a picture signal focused on the CCD 2 through the lens portion 1 is thinned out by 3 in the vertical direction and output as a photographed signal of 1024 x 256 pixels. A digital component signal is supplied from the camera signal processing portion 4 to the memory controller 5. The photographed signal is written to the buffer memory 6 through the memory controller 5. The photographed signal is read at a timing corresponding to a display timing of the LCD 8 and supplied to the D/A converter 7. The D/A converter 7 converts the photographed signal as a digital signal into an analog signal. The analog signal is displayed on the LCD 8. At this point, an area of 960 x 240 pixels is cut from the area of 1024 x 256 pixels written to the buffer memory 6 and the cut area is read from the buffer memory 6 at double speed.

[0029]

Next, the shutter button is pressed and a picture is photographed. In the still picture photographing mode (in the picture format XGA or VGA) as the photographing mode, when the shutter button is pressed, the digital camera is placed in the still picture photographing mode for photographing a still picture. In the still picture photographing mode, the CPU 12 causes the CCD 2 to operate in the photographing

mode. Thus, the CCD 2 outputs a high resolution picture (in the picture format XGA) at a rate of 10 frames per second. Under the control of the memory controller 5, a photographed picture (original picture data (in the picture format XGA or VGA)) is directly stored to the DRAM 9 corresponding to DMA (Direct Memory Access) method.

[0030]

When original picture data is stored to the DRAM 9, under the control of the CPU 12, the original picture data is compressed by the encoder/decoder 15. The compressed picture data (JPEG data) is stored to the DRAM 9. In this case, the JPEG data is stored to an area different from the area of the original picture data. Thereafter, under the control of the CPU 12, the JPEG data is read from the DRAM 9. The JPEG data is written to a particular area of the external storing medium 11 (for example, a floppy disk) through the interface 10.

[0031]

In addition, according to the embodiment of the present invention, a function for recording/reproducing an audio signal corresponding to a photographed still picture or a photographed moving picture is provided. With a trigger of which the shutter button is pressed, audio data is recorded for a predetermined time period. In Fig. 1, reference

numeral 16 is a microphone. An audio signal is supplied from the microphone 16 to an A/D converter 18 through an amplifier 17. The A/D converter 18 samples the audio signal at a frequency of 32 kHz so as to convert the sampled signal as an analog signal into a digital audio signal. The digital audio signal is supplied from the A/D converter 18 to the memory controller 5. The digital audio signal is temporarily stored to a buffer memory of the memory controller 5.

[0032]

The CPU 12 reads the content of the buffer memory by an interrupt process and compresses the digital audio signal in MPEG audio layer2 format (ISO 1172-3) by a software process. The encoding process in the MPEG audio layer 2 format includes a sub-band encoding process, a scaling process, and a bit allocating process. In this case, the encoding process may be performed in MPEG audio layer 1 format or MPEG audio layer 3 format. An MPEG audio stream generated by the software compressing process is written to the DRAM 9. When the MPEG audio stream is written to the DRAM 9, under the control of the CPU 12, a multiplexing process for the MPEG audio stream and the MPEG video stream is performed and the resultant stream is written as a system stream to the DRAM 9. The system stream that is read from the DRAM 9 is recorded to the external storing medium such as a floppy disk in a

general-purpose format through the interface 10 such as a floppy disk controller.

[0033]

In the audio attached moving picture photographing mode, when the shutter button is pressed, the digital camera is placed in a moving picture photographing mode for photographing a moving picture. In the moving picture photographing mode, the CCD 2 operates in the E-to-E mode unlike with the above-described still picture photographing mode. The CCD 2 outputs a photographed signal of which the number of lines is thinned out by 3. This is because in the moving picture photographing mode, since it is necessary to follow the motion of a picture, the amount of picture data should be prevented from increasing. In the moving picture photographing mode, when the shutter button is pressed, pictures are photographed at intervals of a predetermined time period (for example, 5 seconds). However, with the operation of the shutter button, the time period for photographing a moving picture can be prolonged.

[0034]

In the moving picture photographing mode, one of picture format CIF and QCIF is set as a picture size. The memory controller 5 performs a pixel number converting process corresponding to the selected size. The encoder/decoder 15 compresses the picture data

received from the memory controller 5. The compressed picture data (MPEG data) is stored to the DRAM 9. After the picture compressing process and the picture storing process have been completed, as in the still picture photographing mode, under the control of the CPU 12, the MPEG data is written to a predetermined area of the external storing medium 11. For example, in the picture format (picture size) CIF, a moving picture of 15 seconds can be recorded on one floppy disk. In the picture format QCIF, a moving picture of 60 seconds can be recorded on one floppy disk.

[0035]

When a still picture (in the picture format XGA or VGA) is reproduced from the external storing medium 11, JPEG data is read from the external storing medium 11 through the interface 10. The JPEG data is decompressed by the encoder/decoder 15. The decompressed still picture data is written to the DRAM 9. The memory controller 5 reads the still picture data from the DRAM 9 corresponding to the DMA method. The still picture data is transferred to the buffer memory 26 and displayed on the LCD 8. In this case, the number of pixels of the still picture is converted by the memory controller 5. Thus, the reproduced picture is displayed with the same number of pixels as the E-to-E mode.

[0036]

When a moving picture is reproduced from the external storing medium 11, MPEG data (a moving picture file) that is read from a floppy disk is written to the DRAM 9. The data that is read from the DRAM 9 is decompressed in the MPEG format by the encoder/decoder 15. The number of pixels of the decompressed picture data is converted by the memory controller 5 corresponding to the size of the picture that has been recorded. The resultant data is displayed on the LCD 8. When a moving picture (in the picture format CIF or QCIF) is reproduced and displayed, it is displayed in a reduced size on the LCD 8.

[0037]

When a still picture or a moving picture and an audio signal corresponding thereto are reproduced, a system stream reproduced from the external storing medium 11 is stored to the DRAM 9. The CPU 12 separates an audio stream from the system stream that is read from the DRAM 9 and decodes the audio stream in the MPEG audio format. The resultant audio stream is transferred to the buffer memory of the memory controller 5. A D/A converter 19 converts the audio stream as a digital signal into an analog signal. The resultant analog audio signal is reproduced by a speaker 21 through an amplifier 20.

[0038]

According to the embodiment of the present

invention, when a still picture is photographed, original picture data is stored to the DRAM 9. Thereafter, the encoder/decoder 15 compreses the picture data in the JPEG format and stores the resultant data as JPEG data to another area of the DRAM 9. Thereafter, the JPEG data is stored to the external storing medium 11. When a moving picture is photographed, one picture is stored to a working area of the DRAM 9. The picture is compressed by the encoder/decoder 15 in the MPEG1 format. The resultant compressed data as MPEG data is stored to another area of the DRAM 9. This process is performed for each moving picture. The MPEG data is stored to the external storing medium 11. When a moving picture is photographed, an audio attached moving picture photographing operation of which audio is recorded along with a moving picture is performed.

[0039]

In addition to the still picture photographing operation, an audio attached still picture photographing operation can be performed. In other words, when a still picture is photographed, for a predetermined time period after the shutter button is pressed or while the shutter button is being pressed, an audio signal is recorded as an MPEG audio stream. The MPEG audio stream and an MPEG video stream of a still picture are multiplexed as a system stream. The

system stream is written to the DRAM 9 and also recorded to the external storing medium 11.

[0040]

Next, the audio attached still picture photographing operation will be described. Picture data of one picture in high resolution (the picture format XGA or VGA) photographed by the CCD 2 in the photographing mode is stored to the DRAM 9. The original picture data is read from the DRAM 9. The memory controller 5 converts the number of pixels of the picture data and generates a reduced picture in the picture format CIF. The reduced picture is compressed by the encoder/decoder 15 in the MPEG format. An I picture is generated with the original reduced picture. The I picture is written to the DRAM 9.

[0041]

The I picture is followed by a picture whose data amount is fixed (namely, moving vectors of all macro blocks are 0) and that is a predictively encoded picture with the preceding picture (namely, a P picture) or a predictively encoded picture with the preceding picture and the following picture (namely, a B picture). The time period of a P picture or a B picture is almost equal to the time period of the audio signal. When such a video stream is decoded and displayed, the picture of the preceding frame is copied and displayed. Thus, apparently, for the time period

of a P picture or a B picture, a still picture can be displayed.

[0042]

For a predetermined time period after the shutter button is pressed as a trigger (for example, while the shutter button is being pressed), an audio signal is supplied to the buffer memory of the memory controller 5 through the microphone 16, the amplifier 17, and the A/D converter 18. The CPU 12 encodes the audio data stored in the buffer memory corresponding to the MPEG audio format so as to generate an MPEG audio stream.

[0043]

The CPU 12 multiplexes the MPEG video stream and the MPEG audio stream and generates the resultant stream as an MPEG system stream. The MPEG system stream is stored to a record data area of the DRAM 9. The system stream stored in the record data area of the DRAM 9 is recorded to the external storing medium 11 (for example, a floppy disk) through the interface 10.

[0044]

After an MPEG system stream (a multiplexed stream of a video stream and an audio stream) has been recorded to the external storing medium 11, the original picture data (in the picture format XGA or VGA) is read from the DRAM 9. The encoder/decoder 15 compresses the original picture data in the JPEG format

and outputs a JPEG still picture stream. The JPEG still picture stream is rewritten to the record data area of the DRAM 9. The still picture stream stored in the record data area of the DRAM 9 is recorded to the external storing medium 11 (for example, a floppy disk) through the interface 10. Thus, in the audio attached still picture photographing operation, a JPEG file containing only a still picture and an MPEG file containing an I picture (photographed at the same time as the still picture) and audio information are simultaneously generated.

[0045]

Next, with reference to Fig. 3, the MPEG encoding process used in the audio attached still picture photographing operation will be described in detail. A picture signal (in the picture format CIF or QCIF into which the number of pixels of a still picture signal in the picture format XGA or VGA is converted) is input from an input terminal 23 of a video signal processing apparatus to an I picture encoder 24. The I picture encoder 24 converts the input picture signal into an I picture corresponding to the MPEG video format. In addition, an audio signal is input from a microphone 16 (or a line input terminal) to an input terminal 25. The audio signal received from the input terminal 25 is supplied to an MPEG audio encoder 26. The MPEG audio encoder 26 converts the audio signal

into a signal corresponding to the MPEG audio format.

[0046]

A P/B picture generator 27 generates fixed data corresponding to the picture size without performing a motion compensation inter-frame predicting process such as a motion detecting process. Thus, it is not necessary to supply a video signal to the P/B picture generator 27. As described above, the fixed data is a code of which moving vectors of all macro blocks thereof are 0 and that is predicted with the preceding picture. Thus, the fixed data is a picture of the preceding frame. More practically, a picture in the picture format CIF or QCIF is treated as one slice. The first macro block and the last macro block of the slice are skipped. The first macro block and the last macro block are encoded in such a manner that the moving vectors thereof are 0. Although one picture may be divided into a plurality of slices, the header information will increase.

[0047]

Since the number of macro blocks to be skipped is encoded, the data amount of a picture generated by the P/B picture generator 27 varies corresponding to the picture size. In reality, the data amount of a P picture corresponding to the MPEG1 format in the picture format CIF is 28 bytes. The data amount of a P picture corresponding to the MPEG1 format

in the picture format QCIF is 19 bytes. Thus, when the same picture is repeatedly placed in a stream and a decoded picture is displayed apparently as a still picture, with such a P or B picture, the data amount can be remarkably decreased.

[0048]

In Fig. 3, reference numeral 28 is an MPEG system encoder. The MPEG system encoder 28 multiplexes signals received from the I picture encoder 24, the P/B picture generator 27, and the MPEG audio encoder 26 corresponding to the MPEG system format and supplies the multiplexed signal as an MPEG system stream to an output terminal 29. As described above, the MPEG system stream is stored to the DRAM 9. The I picture encoder 24 and the P/B picture generator 27 are contained in the encoder/decoder 15 (shown in Fig. 1). The MPEG audio encoder 26 and the MPEG system encoder 28 are accomplished as software processes of the CPU 12.

[0049]

The structure shown in Fig. 3 can be applied to the audio attached moving picture photographing operation as well as the audio attached still picture photographing operation. In the audio attached moving picture photographing operation, a video signal equivalent to one frame of a photographed moving picture (in the picture format CIF or QCIF into which

the number of pixels of a photographed signal of the CCD 2 is converted) is supplied to the I picture encoder 24. In addition, the P/B picture generator 27 generates fixed data without performing a motion compensation inter-frame predicting process.

[0050]

Fig. 4 shows an example of a frame structure of which a P or B picture received from the P/B picture generator 27 is placed after an I picture received from the I picture encoder 24. Each I picture is followed by two P pictures. The two P pictures are generated by the P/B picture generator 27. The data amount of a P picture is much smaller than that of an I picture. In the example shown in Fig. 3, one out of three frames are thinned out. Thus, the frame rate is 1/3. Consequently, a frame rate of for example 25 Hz that satisfies the minimum frame .rate of the MPEG standard can be accomplished. However, the number of P or B pictures placed between I pictures depends on a desired frame rate. When at least one P or B picture is placed between I pictures, the frame rate can be decreased.

[0051]

Next, with reference to Figs. 5 and 6, an example of the structure (pack structure) of a system stream generated by the MPEG system encoder 28 will be described. Fig. 5 shows a pack structure in the audio attached moving picture photographing operation. Fig.

6 shows a pack structure in the audio attached still picture photographing operation. The pack structure in the moving picture photographing operation is based on a system stream corresponding to the MPEG1 format. In addition, to effectively multiplex data streams, the pack structure has the following features.

[0052]

The size of one pack is fixed. One pack contains audio access units and video access units so that the time period of the audio access units is equal to the time period of the video access units. For example, one pack contains 10 audio frames and 9 video frames. The time period of one video frame is 1/25 seconds. An access unit in the MPEG audio Iayer2 format contains 1152 samples per frame. The audio sampling frequency is 32 kHz. Thus, the time period of the above-described information is equivalent to 0.36 seconds.

[0053]

In addition, one packet contains data of a multiple of access units. In addition, an audio packet with a fixed length is placed at the beginning of a pack. One packet is placed every three video frames (for example, one I picture and two P pictures). A padding stream packet (dummy data) is placed at the end of a pack so that the length of the pack is fixed.

[0054]

As shown in Fig. 5, the first packet contains 10 frames of audio information. Each of the second, third, and third packets contains three frames of video information. The last packet contains a padding stream.

[0055]

In such a pack structure, when a picture is photographed, data that is output from the audio encoder and the video encoder can be multiplexed on real time basis, not buffered. In addition, an I picture is contained in a pack with a fixed length by a rate control. Since the length of the pack is fixed, values of SCR (System Clock Reference) and PTS (Presentation Time Stamp) can be represented with a simple adding process.

[0056]

Next, with reference to Fig. 6, a pack structure in the audio attached still picture photographing operation will be described. Pack 1 (first pack) contains a still picture (I picture). In other words, pack 1 contains an audio packet, a video packet having an I picture of which a still picture (reduced picture) has been encoded, and a P or B picture (at least one picture) of which moving vectors of all macro blocks thereof are 0 and that has been predicted with the preceding frame. Pack 2 contains an audio packet and a P or B picture (at least one

picture).

[0057]

When a picture is encoded, the first pack (pack 1) is encoded so that a still picture and audio can be reproduced on the decoder side. In the later packs, to reduce the data amount, a structure of pack 2 is placed. Thus, while a still picture is being displayed, audio corresponding thereto can be reproduced. Since video information is required for a time period equal to that of audio information to be recorded, video packets for the time period are placed with the structure of pack 2. However, when it is not necessary to reduce the code amount, a system stream may be composed with the structure of pack 1.

[0058]

As an example of the structure of a pack, as with pack 3 shown in Fig. 6, the number of packets per pack may be one rather than the structures of pack 1 and pack 2. As with pack 4 and pack 5, an I picture and a P or B picture may be placed in two successive packs. In addition, there are a plurality of still pictures to be displayed. In this case, when pack 1 is placed in a stream at intervals of a predetermined time period, while different still pictures are being reproduced, audio data corresponding thereto can be reproduced as a slide show.

[0059]

According to the embodiment of the present invention, the encoder/decoder 15 should encode/decode a picture corresponding to the JPEG format and MPEG format. Fig. 7 shows an example of the structure of the encoder/decoder 15. In the embodiment of the present invention, when a picture is encoded corresponding to the MPEG format, an inter-frame motion compensation predictive process is omitted. As a result, a structure that shares the DCT process between the JPEG encoder and the MPEG encoder can be effectively used.

[0060]

In Fig. 7, a picture data as blocks (each of which is composed of (8 x 8 pixels)) is supplied to an input terminal 31. The picture data is supplied from the input terminal 31 to a DCT portion 32. The DCT portion 32 performs a cosine transform process for the picture data and generates 64 coefficients (one DC component and 63 AC components) corresponding to individual pixel data of each block. The coefficient data is supplied to a scanning portion 33. The scanning portion 33 outputs one of two scanning methods (zigzag scanning method and alternate scanning method).

[0061]

An output signal of the scanning portion 33 is supplied to quantizing portions 34a and 34b. The quantizing portions 34a and 34b quantize the

coefficient data using respective scaling factors. One of quantized outputs is selected by a switch circuit SW1. When the JPEG encoding process is performed, the switch circuit SW1 selects the quantized output of the quantizing portion 34a. When the MPEG encoding process is performed, the switch circuit SW1 selects the quantized output of the quantizing portion 34b.

[0062]

The quantized output selected by the switch circuit SW1 is supplied to a JPEG variable length code encoding portion 35a and an MPEG variable length code encoding portion 35b. Since the JPEG variable length code encoding process and the MPEG variable length code encoding process use different Huffman tables each other, two Huffman tables 35a and 35b are provided. When the JPEG encoding process is performed, the AC components of the coefficient data are encoded with variable length code by the variable length code encoding portion 35a and the Huffman table 36a. The encoded output is1 selected by the switch circuit SW2. When the MPEG encoding process is performed, the AC components of the coefficient data are encoded with variable length code by the variable length code encoding portion 35b and the Huffman table 36b. The encoded output is selected by the switch circuit SW2.

[0063]

The switch circuit SW2 is connected to header

adding portions 37a and 37b. The header adding portion 37a adds a header corresponding to the JPEG format to the stream. The header adding portion 37b adds a header corresponding to the MPEG format to the stream. The resultant stream is obtained from an output terminal 38 through a switch SW3 that operates corresponding to whether the JPEG encoding process or the MPEG encoding process is performed.

[0064]

Although the quantizing portions 34a and 34b are shown as different structural elements, many parts of them can be structured in common as hardware. Likewise, many parts of the header adding portions 37a and 37b, the JPEG variable length encoding portion 35a, and the MPEG variable length code encoding portion 35b can be structured in common as hardware. On the other hand, the Huffman tables 36a and 36b should be separately provided as hardware. Fig. 7 shows the structure of the encoder portion of the encoder/decoder 15. The decoder portion is composed of a header separating portion, a variable length code decoding portion, an inversely quantizing portion, and an inversely DCT portion. As with the encoder portion, many portions of the decoder portion can be structured in common as hardware. Since the inter-frame motion compensation predictive process is omitted from the MPEG format encoding process, the hardware scale of the

encoder/decoder can be decreased. Thus, an integrating circuit of the encoder/decoder can be easily designed.

[0065]

According to the present invention, as examples of the external storing medium 11, various types of disk mediums such as a detachable card and a floppy disk can be used. In addition, the encoding process according to the present invention can be applied for data transmissions to a network, RS232C, non-contact type IrDr, and so forth.

[0066]

[Effect of the Invention]

According to the invention of claim 1, when still picture data is recorded or transmitted, audio information corresponding thereto can be multiplexed with the still picture data. Thus, a still picture and audio information that has been recorded can be reproduced by a personal computer using general purpose software.

[0067]

According to the invention of claim 4, when a still picture is photographed, audio information corresponding thereto can be recorded. The still picture and the audio information can be multiplexed corresponding to the MPEG format. Thus, a still picture and audio information that has been recorded can be reproduced by a personal computer using software

that is commercially available.

[0068]

According to the invention of claim 9, a function for simultaneously recording a still picture and an audio signal can be accomplished for a digital camera. In addition, when an audio attached still picture is recorded, only a still picture can be recorded. Thus, corresponding to a desired application, recorded data can be used.

[Brief Description of the Drawings]

[Fig. 1]

Block diagram showing the overall structure of a digital camera according to an embodiment of the present invention.

[Fig. 2]

Schematic diagram for explaining a picture size according to an embodiment of the present invention.

[Fig. 3]

Block diagram showing an example of the structure of an encoding apparatus according to the present invention.

[Fig. 4]

Schematic diagram showing a frame structure of an output signal of the encoding apparatus according to the present invention.

[Fig. 5]

Schematic diagram showing an example of a data structure of a system stream that is output from an encoding apparatus in an audio attached moving picture photographing operation.

[Fig. 6]

Schematic diagram showing an example of a data structure of a system stream that is output from an encoding apparatus in an audio attached still picture photographing operation.

[Fig. 7]

Block diagram showing the structure of an encoder/decoder according to an embodiment of the present invention.

[Description of Reference Numerals]

2 ... CCD, 4 ... Camera signal processing portion, 5 ... Memory controller, 8 ... LCD, 9 ... DRAM, 11 ... External storing medium, 12 ... CPU, 13 Operation and inputting portion, 15 ... Encoder/decoder

[Title of Document] Abstract

[Abstract]

[Subject]

To generate an output of audio attached photographed still picture corresponding to MPEG format.

[Solving means]

A still picture is photographed by a CCD 2. At the same time, an audio signal received from a microphone 16 is recorded. The still picture and the audio data are written to a DRAM 9 through a remote controller 5. The still picture data is supplied to an encoder/decoder 15. The encoder/decoder 15 compresses the still picture corresponding to MPEG video format. Software causes a CPU 12 to compress the audio data corresponding to MPEG audio format. The compressed video data and the compressed audio data are multiplexed and stored to a DRAM 9. When the still picture is compressed corresponding to the MPEG video format, data of one picture is encoded. Thus, an I picture is generated. In addition, a P picture of which all macro blocks thereof are zero and the chronologically preceding picture is copied as an encoded picture is generated. Thus, an output of a frame structure of which an I picture is followed by at least one P picture is generated. The multiplexed data is recorded to an external storing medium 11.

[Selected Drawing] Fig. 1